UNIVERSITY OF COPENHAGEN

MASTER THESIS

Deep hedging

Author:
Peter Pommergård LIND

Supervisor:
Dr. David SKOVMAND

A thesis submitted in fulfillment of the requirements
for the degree of Master Thesis in Actuarial Mathematics

August 6, 2020

Declaration of Authorship

I, Peter Pommergård LIND, declare that this thesis titled, "Deep hedging" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

"You were hired because you met expectations, you will be promoted if you can exceed them."

Saji Ijiyemi

UNIVERSITY OF COPENHAGEN

Abstract

Department of Mathematical Science
Science

Master Thesis in Actuarial Mathematics

Deep hedging

by Peter Pommergård LIND

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Acknowledgements

The acknowledgments and the people to thank go here, don't forget to include your project advisor...

Contents

List of Figures

List of Tables

List of Abbreviations

B-S	Black-Scholes
BM	Brownian Motion
FPT1	Fundamental Pricing Theorem I
FPT2	Fundamental Pricing Theorem II
GBM	Geometric Brownian Motion
LIBOR	London Interbank Offered Rate
SDE	Stochastic Differential Equation
S-F	Self-Financing

List of Symbols

c	European call option price
p	European put option price
S_0	Stock price today
K	Strike price
T	Maturity in years
σ	Volatility of stock price
C	American Call option price
P	American Put option price
S_T	Stock price at option maturity
r	Continuous compounding risk-free rate
$V^h(t)$	Value process
X	Simple Derivative
Φ	Contract function
W_t	Weiner process (a synonym brownian motion).

For/Dedicated to/To my...

Chapter 1

Introduction

In recent years we have seen an increasing complexity of financial products, where big investment- and banks use a lot of money on financial engineerers in creating new innovative products. With the complexity a lot of challenges has risen in this field. Nevertheless the products can help to mitigate risk and leverage your portfolio. A recent example from the financial crisis in 2007 where credit default swap (CDS) almost led to AIGs bailout. A CDS is a derivative, where you insure your risk of losing money on some financial product. The strategy of writing CDS seemed like a good business for AIG as long there was a bull market, because they got good feeds for insuring credit. The CDS was the main reasions that AIG needed a bailout by the US government under the recent financial crisis. In hindsight they wrote to many CDS, hence AIG was too exposed for risk. A great understanding in the financial derivatives is important to understand your risks and ulitmately migate the damage of financial turmoil as Warren Buffett says derivatives is "Financial weapons of mass destruction" (page 15 (Buffett, 2002)). Eventhough Buffett is critical against derivatives he acknowlegde the usage of derivatives, because he owns derivatives in his portfolio. Derivatives gives the trader more options either to utilize arbitrage, speculate or hedge, but without care or knowledge about your book of derivative the outcome can be disastrous (Buffett, 2008).

The focus is on financial derivatives, where the prime examples will be plain vanilla stock options. We will start with the most basic derivatives European options and move toward more complex products like American options. The European option will be the reference point for our different nummerical approaches to American options, because the European option has a closed form solution (see proposition 2.2.1). When moving into more complex derivatives as American options the Black Scholes analytical framework breaks down, and this calls for numerical methods. We will take different numerical approaches for pricing and hedging, where the ultimate goal is to use machine learning for pricing and hedging.

Chapter 2

Arbitrage theory in continuous time finance
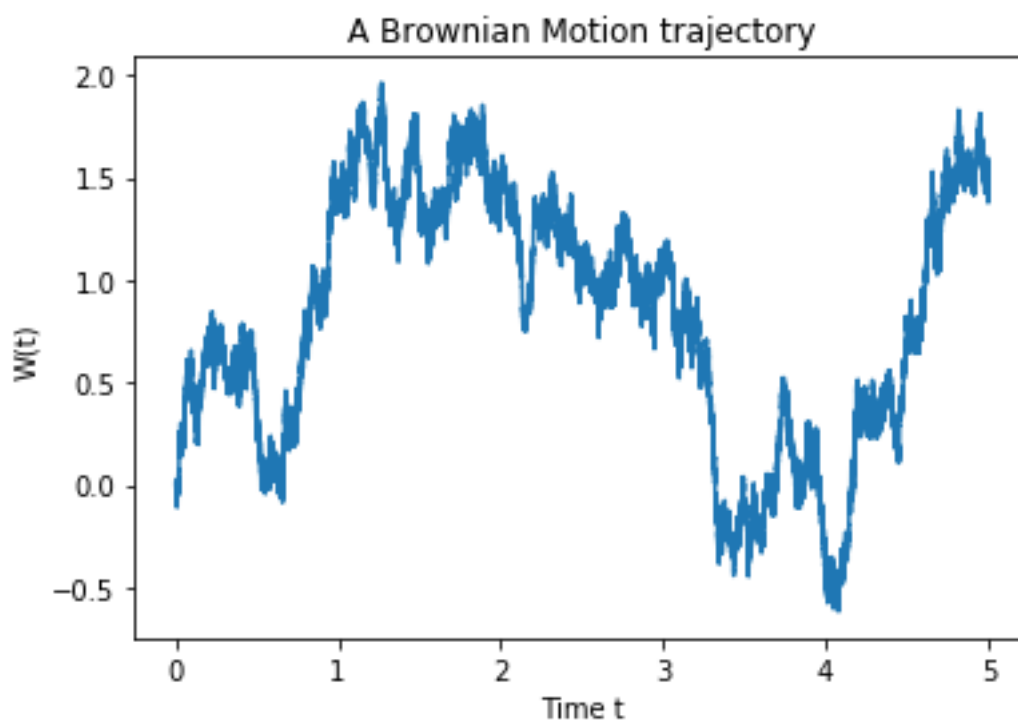
2.1 Financial markets

FIGURE 2.1

In the financial markets there is a lot of players and different types of investments. The classical investments are bonds and stocks, where you either lending or buying equity. The big players are commercial banks, investment banks, insurance companies and pension funds. The derivatives are depending on an underlying asset, where the dependency is specified in the contract. The options discussion in the introduction are all depending on a underlying stock. The contract can be constructed in many ways, hence it gives more options to construct your portfolio (see Appendix A for examples). When pricing financial product we use the market to price derivatives (This correspond to the equivalent martingale measure $\mathbb{Q}$ to the objective measure $\mathbb{P}$), so we do not introduce arbitrage to the market. In the classic

Black Scholes formula for European options, we will assume following about the market:

Assumption 2.1.1. *We assume following institutional facts:*

- *Short positions and fractional holding are allowed*

- *There are no bid-ask spread, i.e. selling price is equal to buying price*

- *There are no transactions costs of trading.*

- *The market is completely liquid, i.e. it is possible to buy/sell unlimited quantities on the market. You can borrow unlimited amount from the bank by selling short.*

(see p. 6 (Björk, 2009))

We can discuss these assumptions at length, but in order to progress mathematically, we need to accept them for now. There is some justification for liquidity on vanilla options, because those options gets traded on large scale. Before going into the mathematics of the Black Scholes formula, we need to introduce key concepts.

2.1.1 Financial Derivatives

There a broad range of different derivatives. In this thesis, we will mainly divide derivatives into two classes.

1. Simple derivatives (T-claims)

2. Exotic derivatives

The first class is the simple derivaties or T-claims. These are simple because you can only exercise them at maturity (time T). The exotic derivatives is all kind of functions on the underlying assets, where you have more options than exercise at termination time. There are so many derivatives, hence the list will not be comprehensive at all. Some important simple derivatives will be the European calls and puts, because we can price analytically.

Definition 2.1.1. European Call Option: A Europeann call option is a option where the owner of the option has the option to exercise at maturity. The contract function for the derivative:

$$\phi(S(T)) = \max\{S(T) - K, 0\} \tag{2.1}$$

Where S(T) is the price of underlying asset at maturity and K is the agreed strike price.

For illustration of above contract see appendix A.
(Björk, 2009)

2.1.2 Self-financing portfolio (Without consumption)

A self-financing portfolio h, is a portfolio h which doesn't get any external injection of money. h is the number of each assets in our portfolio. We denote $V^h(t)$ the value of our portfolio h at time t, hence:

Definition 2.1.2. Self-financing portfolio A portfolio consisting of N+1 assets: h(t)=$(h_0(t), h_1(t), \ldots, h_N)$ is self-financing if:

$$dV^h(t) = \sum_{i=0}^{N} h_i(t)dS_i(t) \tag{2.2}$$

Where S_i is the $i'th$ asset in our portfolio, N+1 is the total number of assets and $V^h(t) = \sum_{i=0}^{N} h_i(t)S_i(t)$

The important takeaway is that a S-F portfolio is kind of a budget restriction. You are only allowed to reallocate your assets within the portfolio but not injecting cash into the portfolio. The concept is important for the discussion of arbitrage and hedging.

2.1.3 Arbitrage

Arbitrage is the financial term for a "free lunch". An investor can profit without bearing risk, if there is arbitrage on the market. In order to avoid making a "money machine", we want to price derivatives to be arbitrage free.

Definition 2.1.3. Arbitrage: An arbitrage possibility on a financial market is a self-financed portfolio h suct that

$$\begin{aligned} V^h(0) &= 0 \\ P(V^h(T) \geq 0) &= 1 \\ P(V^h(T) > 0) &> 1 \end{aligned} \tag{2.3}$$

We say that the market is arbitrage free if there aro no arbitrage possibilities.
(see p. 96 (Björk, 2009))

From the definition a self-financing portfolio fullfilling equation (2.3) would give the possiblility for arbitrage. The investor in this portfolio starts with 0 dollars, and without injecting any money, the investor is certain of not losing any money. In addition he has a positive probability by ending up with more than 0 at maturity. Arbitrage is a way to price financial products "fair". To price "fair" and hedge againt risk will be the topics for this thesis.

2.1.4 Complete Market and Hedging

Hedging is a concepts to protect against exposure to risk. A hedge is simply a risk neutralization action in order to minimize the overall risk. In the defition below, we define a hedge for an simply T-claim (**??**).

Definition 2.1.4. Hedging and completeness for T-claim: A T-claim X can be hedged, if there exist a self-financing portfolio h s.t.:

- $V^h(T) = X$ P-a.s.

I.e. h is an hedge portfolio for X if it is guaranteed to pay in all circumstances an amount identical to the payout of X.
The market is complete, if every derivative is hedgable. (see p. 115 (Björk, 2009))

Hedging and completeness means the same for other derivatives than T-claims, but for now we will only show the concepts for the T-claim.

2.2　Black-Scholes Formula two dimensionel

In addition to our assumptions for the financial market, we also assume:

Assumption 2.2.1. *Black-Scholes assumptions We assume following ideal conditions in addition to* (2.1.1)*:*

- *The short-term interest rate is known and is constant through time*

- *The stock price follows a Geometric Brownian Motion. The σ is constant.*

- *The stock pays no dividends or other distributions.*

- *The option is a simple option ("European" see* (2.1.1)*).*

(see p. 640 (Black and Scholes, 1973))

We assume the underlying stock follows a geometric brownian motion: $dS(t) = \alpha S dt + \sigma S dW_t$ where the solution to the SDE is given as

$$S(t) = S(0) \cdot \exp\left((\alpha - \frac{1}{2}\sigma^2)t + \sigma W(t) \right) \qquad (2.4)$$

The μ and σ have clear empirical meanings.

Theorem 2.2.1. *Black-Scholes PDE: blabla*

$$F_t(t,s) + rsF_s(t,s) + \frac{1}{2}s^2\sigma^2(t,s)F_{ss}(t,s) - rF(t,s) = 0 \qquad (2.5)$$

$$F(T,s) = \Phi(s) \qquad (2.6)$$

The below proposition is a consequence of the B-S equation:

Proposition 2.2.1. *Black-Scholes formula for call option: The price of a European call option with strike K and maturity T is given by the formula $\Pi(t) = F(t, S(t)$, where*

$$F(t,s) = s \cdot N(d_1(t,s)) - e^{-r(T-t)} \cdot K \cdot N(d_2(t,s))$$

N is the cumulative distribution function of a standard normal distribution $\mathcal{N}(0,1)$ and

$$d_1(t,s) = \frac{1}{\sigma \cdot \sqrt{T-t}} \cdot \left(\ln(\frac{s}{K}) + (r + \frac{1}{w}\sigma^2)(T-t) \right)$$
$$d_2(t,s) = d_1(s,t) - \sigma\sqrt{T-t}$$

(see p. 105 (Björk, 2009))

The above formula for the European call option is actually the same for an American call option, but is not true for an American put option or for call options paying dividends. The result for the American call option was shown by Merton (Merton, 1973), that the intrinsic value is never greater than the worth of the option given by the risk-neutral valuation formula (Björk, 2009).

Theorem 2.2.2. *Risk-neutral valuation formula: Given Q is the EMM*

$$\Pi(t, X) = exp(-r(T-t)) \cdot E_{t,x}^Q[X] \qquad (2.7)$$

Proposition 2.2.2. *Put-call parity: Assume the call and put option has same strike price and time to maturity.*

$$p(t,s) = K \cdot \exp(-r(T-t)) + c(t,s) - s$$

Chapter 3

Classical numerical results

In this section will we look at the two classical valuaing algorithms in computational finance the Binomial model and the Least Square Monte Carlo (LSM) approach. The models will both serve as reference for the Machine Learning model and provides insight into valuation of American options.

3.1 Binomial Pricing model

The Binomial model provides an intuitive and easy implementable model for valuing American and European options. The Binomial model comes handy, when no analytical model exists for American options. The Binomial model also has its limitations, because it is not suited for valuing path dependent options or options with several underlying factors. The key difference on the Binomial model and the other numerical procudures is that the Binomial model is build on a discrete framework.

The central concepts arbitrage and completeness from continuous time also work in the discrete time setup. The paper (Cox and Stephen Ross, 1979) which introduced the binomial model to option pricing came after the Black-Scholes model described in section 2 (Black and Scholes, 1973). The main reason for developing a model in discrete time, is that the the discrete time approach gives a simplified model in terms of the mathematics and highlights the essential concepts in option pricing theory. You can argue that the simpler mathematics in this model makes the binomial model more instructive and clear. Besides being easier to understand for non-mathematician it works nicely with other options than the European options like American options.

Eventhough we assume the stock price moves at discrete time instead in continuous time. It can actually be shown for a European Option that if the number of timesteps in the tree aproaches infinity, then the binomial model will converge to the continuous time closed form solution for a European option (Cox and Stephen Ross, 1979) (Hull, 2018). Hence the binomial pricing model will be equivalent with the continuos time analytical pricing model derived by Fischer Black and Myron Scholes in the limit for european options (Cox and Stephen Ross, 1979).

To value a american put option, we lay out all the possible path of the stock, based on the S_0, σ and T. We need to specify the number of timesteps ($\Delta t = \frac{T}{N}$ where $N = No. of steps$) for the tree, where for each step, we add another possible value for the stock. We only add 1 more possibility for each timestep because the tree recombines. The precision for the algorithme increases with the number of steps and the option value stabilizes (see Figure 3.1). For valuing an american put option, we value the exercise value at maturity (time T) for all possible outcomes for the stock. Then we

work backward in the tree by comparing intrinsic value with the conditional expectation, where we choose the maximum of these two (Hull, 2018).
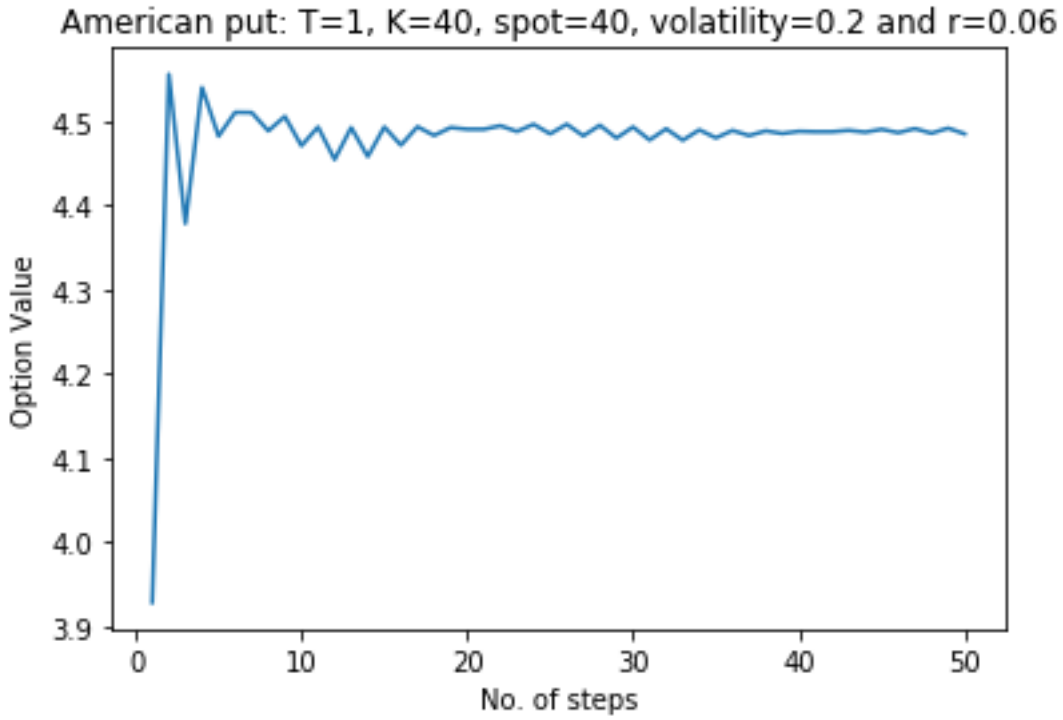

American put: T=1, K=40, spot=40, volatility=0.2 and r=0.06

FIGURE 3.1

3.1.1 Mathematics in Binomial valuation model

The mathematics behind the binomial model is simpel and we will in this section provide the basic mathematics. First we need to construct the tree, then afterwards work backwards in the tree for valuation. For each time step (Δt), we assume the stock (S) can move up (u) or down (d). In order to avoid arbitrage we find the risk neutral measure q for the binomial tree, where q is the probability for the stock moves up. The risk neutral measure q is chosen s.t. the expected return is the risk-free rate r.

Theorem 3.1.1. *Risk-neutral valuation formula in discrete time. Assume there exists a risk free asset. Then the market is arbitrage free if and only if there exists a risk neutral measure $Q \sim P$ s.t.*

$$s = \exp(-r\Delta t) \cdot E^Q[S(t + \Delta t)|S(t) = s] \tag{3.1}$$

Where Δt is a single timestep.

From the above theorem, we can calculate the risk neutral mesure as:

$$q = \frac{e^{\Delta t} - d}{u - d}$$

The d and u is chosen s.t. they match volatility. So we choose:

$$u = \exp(\sigma\sqrt{\Delta t}) \quad d = \exp(-\sigma\sqrt{\Delta t})$$

Now we have determined the three parameters needed for constructing a binomial tree (Cox and Stephen Ross, 1979) (Hull, 2018) (Björk, 2009).

We want to value an American put option, hence we need to work backward in the tree and comparing in each node the intrinsic value with the conditional expection (see theorem 3.1.1) by:

$$max\{K - S(t), \exp(-r\Delta t) \cdot E^Q[P(t + \Delta t, T)|P(t, T) = p]\} \tag{3.2}$$

The comparision will be applied for every node in each timestep Δt and all the way back in time to the initialization date. By this precedure we get present value of the American option at initialization.

3.2 Least Square Monte Carlo Method

The other classical result in this section is of an different nature, because it is based on simulation and linear regresion. In our setting we regress the expected payoff by continuation of the contract and compare it to the intrinsic value. The dependent variable in the regression is the expected value of continuation and the independent variables is a set of othogonal basis functions in $L^2(\Omega, \mathcal{F}, Q)$ of the simulated paths. Typical choices for basis functions could be weighted Laguerre -, Hermit -, and Jacob polynomials. This kind of regression is a nonlinear expansion of the linear model. In order to create data, we will simulate paths according to the underlying risky asset.

3.2.1 LSM method for an American put

We want to valuate an American put option with a stock as underlying asset. We take the same assumptions as in Chapter 2 (see assumption 2.2.1) except the option is an American option. Hence in order to simulate the paths of the stock, we simulate from an GBM: $dS(t) = rSdt + \sigma SdW_t$ where σ and r is constant (see solution to SDE equation 2.4). We simulate 100.000 paths for the stock. Like in the binomial model, we work backward to decide the optimal stopping time. The computer is discrete, hence we simulate the stock path as an Bermuda option, where we have 50 timesteps per year.

At maturity the cashflow from the option is the same as for an European put option, hence the cashflow from each path is $C(\omega, T; T, T) = max(K - S_T, 0)$. We use the notation $C(\omega, s; t, T)$ denote the path of cash flows generated by the option condition on the option not being exercised before t and the option holder follw the optimal stopping strategy for all s, $t < s \leq T$. (inspired by (Longstaff and Schwartz, 2001) p. 121). The continuation value is given by:

$$F(\omega; t_k) = E^Q[\sum_{j=k+1}^{K} \exp(-\int_{t_k}^{t_j} r(\omega, s)ds)C(\omega, t_j; t_k, T)|\mathcal{F}_{t_k}] \tag{3.3}$$

where $r(\omega, t)$ is risk free interest rate, and the $\mathcal{F}_{t_k}$ is the filtration at time t_k.

The optimal stopping strategy is then by comparing this continuation value with the intrinsic value at each time step. By working backward in time until the initialization of the option, we have specified the optimal stopping times and the cashflows associated with exercising at the optimal stopping times. To estimate the condition

expection in equation 3.3, we regres with the basis functions taking on the underlying asset for the option being the independent variable:

$$F(\omega; t_{K-1}) = \sum_{j=0}^{\infty} a_j L_j(X)$$

where a is the coefficients for the regression, L is the basis function, where the argument is the underlying asset X (Longstaff and Schwartz, 2001).

3.2.2 Numerical results

By the above two algorithms for valuation, we choose to vary spot, volatility and maturity for pricing an American put option with K=40 and r=0.06. This table will serve as reference for the machine learning algorithm in chapter (!TODO chapter for machine learning). For the binomial tree we use 100 timesteps, which gives stable results (compare to figure 3.1) and for the lsm we use 10^5 paths with 50 timesteps per year. The European option is valued by using BS closed form solution for a call option (see proposition 2.2.1) and Put-call parity (see proposion 2.2.2). We see the

TABLE 3.1: Valuation of American put option with K=40 and r=0.06.

Spot	σ	T	Closed form European	Binomial Tree	LSM	abs. diff.
36	0.2	1	3.844	4.488	4.478	0.010
36	0.2	2	3.763	4.846	4.828	0.018
36	0.4	1	6.711	7.119	7.092	0.027
36	0.4	2	7.700	8.508	8.500	0.008
38	0.2	1	2.852	3.260	3.245	0.015
38	0.2	2	2.991	3.748	3.735	0.013
38	0.4	1	5.834	6.165	6.144	0.021
38	0.4	2	6.979	7.689	7.665	0.024
40	0.2	1	2.066	2.316	2.313	0.003
40	0.2	2	2.356	2.885	2.881	0.004
40	0.4	1	5.060	5.310	5.326	0.016
40	0.4	2	6.326	6.914	6.908	0.006
42	0.2	1	1.465	1.622	1.622	0.000
42	0.2	2	1.841	2.217	2.212	0.005
42	0.4	1	4.379	4.602	4.596	0.006
42	0.4	2	5.736	6.264	6.243	0.021
44	0.2	1	1.017	1.117	1.113	0.004
44	0.2	2	1.429	1.697	1.688	0.009
44	0.4	1	3.783	3.956	3.962	0.006
44	0.4	2	5.202	5.656	5.649	0.007

maximum difference between the two algorithms is 0.027 at S=38, $\sigma = 0.4$ and T=2. The other obvious fact is that the european put has a lower value than its American counterpart, because the continuous exercise feature adds additional value to the put option.

Chapter 4

Deep Learning

4.1 Main concepts in Deep learning

Deep Learning is a specialized field in Machine Learning, which means it is within the field of applied statistics. Like in statistics and Machine Learning the basic components of a Deep Learning algorithm are a dataset, cost function, optimazation algorithm and a model. E.g. in the lsm method we assumed the model was Gaussian, dataset was the simulated paths, the lose function was the mean square error and the optimazation algorithm was solving the normal equations. Deep learning is about studying neural networks which allows for greater flexiblity than standard methods like linear regression. "Deep" comes from that a neural network consists of multiple layers, where the depth tells you how many layers the network has. The network consists of a input layer, hidden layers and finally a output layer where the input and output layers are observable to the user. To fit a model we need to provide weights, bias and the activation function for each layer, hence the ouput is a chain of functions applied to the input. In order to measure the performance of the model, we need a function to measure the difference on the response variable and the actually observed response. This function is referred as the loss function, where the cost function is the average over the loss functions. The cost function tells us how good our model is on the data. The cost function is key to improving our model or in machine learning lingo learning the model, hence we applied a optimization algorithm to find the optimal set of weights in order to improve the cost function. (Goodfellow, Bengio, and Courville, 2016)

4.1.1 Subsection 1

4.1.2 Subsection 2

4.2 Main Section 2

Appendix A

Option contracts

This list of option contracts are far from complete, but the purpose is to illustate some payoff contracts for reference.

A.1 European Call and Put

The European options will be the most basic options, we will work with. This means not that they are not important, actually they are key for pricing options. The European call option is a contract, which pays at maturity $\Phi(S(T)) = max(S - K, 0)$.

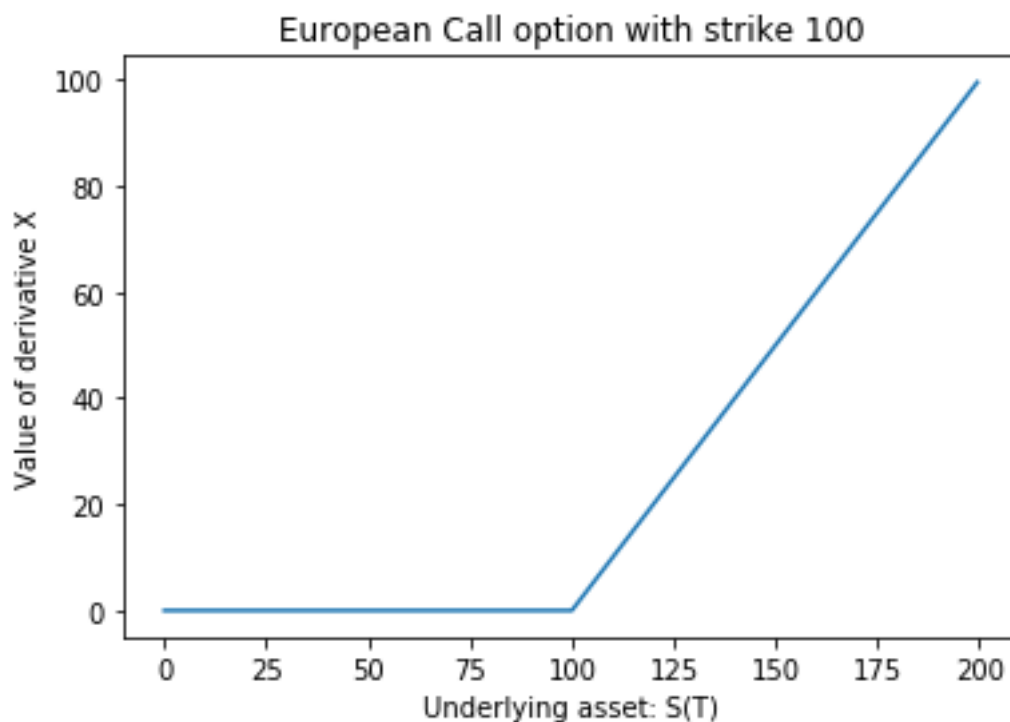
FIGURE A.1: European call with K=100

The European put is very similar to the call, except now we earn, when the stock is below the strike price K.
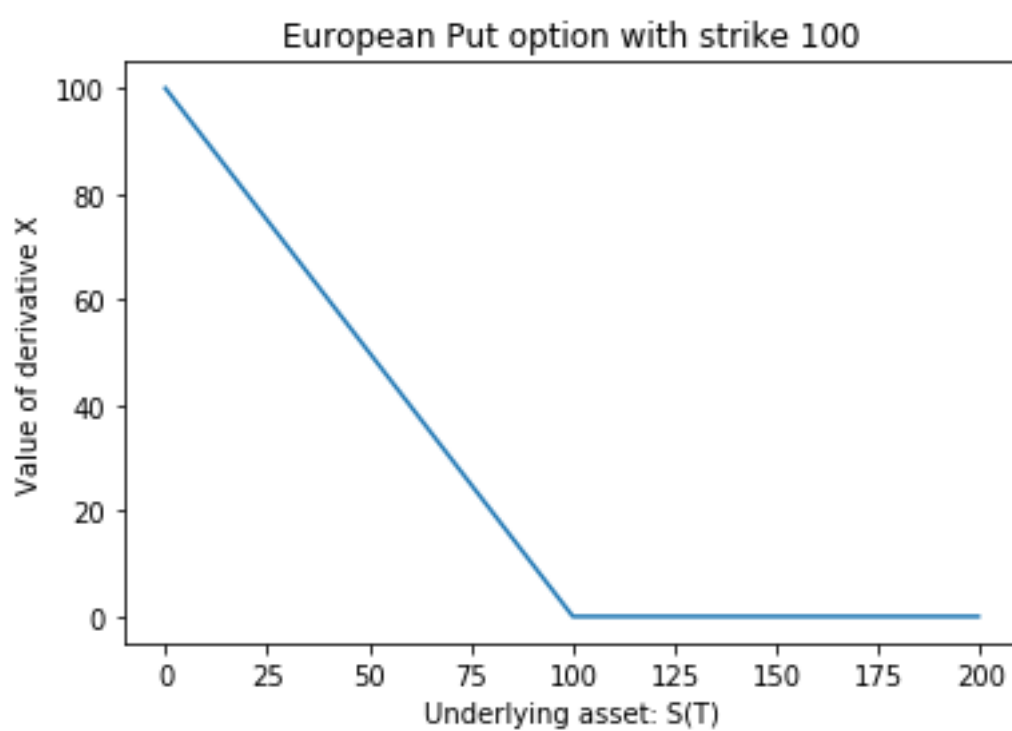$\Phi(S(T)) = max(K - S, 0)$.

FIGURE A.2: European put with K=100

Appendix B

Mathematical definitions

Definition B.0.1. Orthogonal vectors: Two vectors $\vec{a}$ and $\vec{b}$ are orthogonal, if their dot product is 0:

$$\vec{a} \cdot \vec{b} = 0$$

We will use the notation:

$$\vec{a} \perp \vec{b} \tag{B.1}$$

Bibliography

Björk, Thomas (2009). *Arbitrage Theory in Continuous Time*. Third edition. Oxford.

Black, Fischer and Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". In: *The Journal of Political Economy* 81.3, pp. 637–654. URL: http://www.jstor.org/stable/1831029 (visited on 02/09/2020).

Buffett, Warren (2002). "Berkshire Hathaway's (BRK.B) annual letters to shareholders". In: URL: https://www.berkshirehathaway.com/letters/2002pdf.pdf (visited on 06/23/2020).

— (2008). "Berkshire Hathaway's (BRK.B) annual letters to shareholders". In: URL: https://www.berkshirehathaway.com/letters/2008ltr.pdf (visited on 06/23/2020).

Cox, John and Mark Rubinstein Stephen Ross (1979). "Option princing: A simplified approach". In: *Journal of Financial Economics* 7, pp. 229–263.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. http://www.deeplearningbook.org. MIT Press.

Hull, John C. (2018). *Options, Futures, and Other Derivatives*. Vol. Tenth edition. Pearson Education.

Longstaff, Francis A. and Eduardo S. Schwartz (2001). "Valuing American Options by Simulation: A Simple Least-Squares Approach". In: *The Review of Financial Studies*.

Merton, Robert C. (1973). "Theory of Rational Option Pricing". In: *The Bell Journal of Economics and Management Science* 4.1, pp. 141–183. URL: http://links.jstor.org/sici?sici=0005-8556%2819732 1%294%3A1%3C141%3ATOROP%3E2.0.CO%3B2-0 (visited on 05/14/2020).